# Heteroskedasticity

May 19, 2021

Eric B. Rasmusen

*Abstract*

We can correct for heteroskedasticity even when estimating a mean.

Professor of Business Economics and Public Policy, Kelley School of Business, Indiana University, 1309 East Tenth Street, Bloomington, Indiana, USA 47405, (812) 855-9219 (secretary), Erasmuse@indiana.edu, http://www.rasmusen.org/papers/heteroskedasticity.

**1. Introduction** asdfadfds

Serial correlation is about having a random sample. Heteroskedasticity is about having data of different quality.

**2. The Weighted Mean**

We want to estimate the mean $\mu$ from $N$ observations $y_i = \mu + \epsilon_i$, where the $\epsilon_i$ are independently distributed with mean 0 and variance $\sigma_i^2$. Suppose we estimate the population mean by using a weighted mean:

$$\hat{\mu}^* \equiv \frac{\Sigma_{i=1}^N w_i y_i}{N}, \tag{1}$$

where we choose $w_i$ so that $\Sigma_{i=1}^N w_i = N$. We could use $w_i = 1$, for example, which is the same as the unweighted mean.

This is an unbiased estimator.

$$E\hat{\mu}^* = E\frac{\Sigma_{i=1}^N w_i y_i}{N} = E\frac{\Sigma_{i=1}^N (w_i\mu + w_i\epsilon_i)}{N} = \frac{\mu}{N}\Sigma_{i=1}^N w_i + E\frac{\Sigma_{i=1}^N w_i\epsilon_i}{N} = \frac{\mu}{N}N + 0 = \mu \tag{2}$$

The variance of the estimator is

$$\begin{aligned} Var(\hat{\mu}^*) &= E\frac{\left(\frac{\Sigma_{i=1}^N w_i y_i}{N} - \mu\right)^2}{N} \\ &= E\frac{\left(\frac{\Sigma_{i=1}^N w_i\mu + w_i\epsilon_i}{N} - \mu\right)^2}{N} \\ &= E\frac{\left(\mu + \frac{\Sigma_{i=1}^N w_i\epsilon_i}{N} - \mu\right)^2}{N} \\ &= \frac{\frac{\Sigma_{i=1}^N w_i^2\sigma_i^2}{N}}{N} \\ &= \frac{\frac{\Sigma_{i=1}^N w_i^2\overline{\sigma}^2}{N}}{N} \end{aligned} \tag{3}$$

Notice, in particular, that if $w_i = 1$ for all $i$, then

$$Var(\hat{\mu}^*) = \frac{\frac{\Sigma_{i=1}^N \overline{\sigma}^2}{N}}{N} = \frac{\overline{\sigma}^2}{N}. \tag{4}$$

But we can do better than that. Let's choose $w_i$ to minimize the variance. How do we do that? With a giant lagrangian? We want to minimize the variance such that $\Sigma w_i = N$. So we solve

$$Minimize_{w_i} \frac{\Sigma_{i=1}^N w_i^2\sigma_i^2}{N^2} - \lambda(\Sigma w_i - N) \tag{5}$$

This has the first order conditions for each $i$,

$$\frac{2w_i\sigma_i^2}{N^2} - \lambda = 0, \tag{6}$$

so

$$w_i = \frac{N^2\lambda}{2\sigma_i^2} \tag{7}$$

or

$$\lambda = \frac{2\sigma_i^2 w_i}{N^2} = \frac{2\sigma_j^2 w_j}{N^2} \tag{8}$$

Thus, $\sigma_i^2 w_i = \sigma_j^2 w_j$ and $\frac{w_i}{w_j} = \frac{\sigma_j^2}{\sigma_i^2}$. A weight that satisfied that first order condition and the constraint is:

$$w_i^* = \frac{\overline{\sigma}^2}{\sigma_i^2} \tag{9}$$

This puts more weight on the better data, the observations with lower variance.

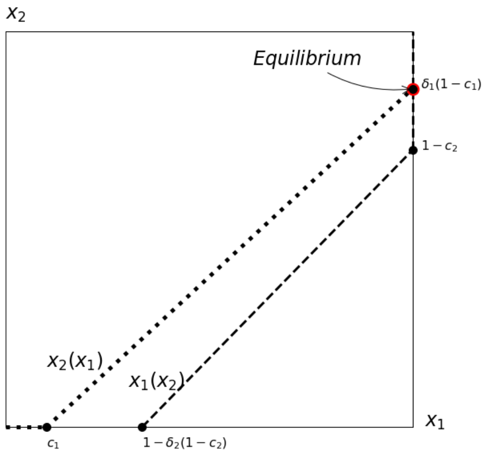
The average variance of the residuals, the sample variance, is (with the N-1 correction so as to be unbiased):

$$\begin{aligned}
Var(\hat{\epsilon}_i{}^*) &= E\frac{\Sigma_{i=1}^N \left(\frac{\Sigma_{j=1}^N w_j y_j}{N} - y_i\right)^2}{N-1} \\
&= E\frac{\Sigma_{i=1}^N \left(\frac{\Sigma_{j=1}^N \left(w_j \mu + w_j \epsilon_j\right)}{N} - \mu - \epsilon_i\right)^2}{N-1} \\
&= E\frac{\Sigma_{i=1}^N \left(\frac{\Sigma_{j=1}^N w_j \epsilon_j}{N} - \epsilon_i\right)^2}{N-1} \\
&= E\Sigma_{i=1}^N \left(\frac{\left(\Sigma_{j=1}^N w_j \epsilon_j\right)^2}{N^2(N-1)} - 2\frac{\Sigma_{j=1}^N w_j \epsilon_j \epsilon_i}{N(N-1)} + \frac{\epsilon_i^2}{N-1}\right) \\
&= \Sigma_{i=1}^N \left(\frac{\Sigma_{j=1}^N w_j^2 \sigma_j^2}{N^2(N-1)} - \frac{2 w_i \sigma_i^2}{N(N-1)} + \frac{\sigma_i^2}{N-1}\right) \\
&= \Sigma_{i=1}^N \left(\frac{\Sigma_{j=1}^N w_j^2 \sigma_j^2}{N^2(N-1)} - \frac{2N w_i \sigma_i^2}{N^2(N-1)} + \frac{N^2 \sigma_i^2}{N^2(N-1)}\right) \\
&= \frac{N^2 \overline{\sigma}^2}{N^2(N-1)} - \frac{2N^2 \overline{\sigma}^2}{N^2(N-1)} + \frac{N^3 \overline{\sigma}^2}{N^2(N-1)} sdfsfsdd \\
&= -\frac{\overline{\sigma}^2}{N-1} + \frac{N\overline{\sigma}^2}{N-1} sdfsdfd \\
&= \frac{(N-1)\overline{\sigma}^2}{N-1} sdfsdf \\
&= \overline{\sigma}^2 sdfdf
\end{aligned} \tag{10}$$

How do we estimate $\sigma_i^2$?

**4. Template stuff**

FIGURE 1:
SDFSAFSDFADFD REACTION CURVES $x_1$ AND $x_2$




## References

Kennedy

White